Best-Partition Multiparty Communication Complexity

Notes by Toniann Pitassi

Abstract

The "number-on-forehead" (NOF) model of multi-party communication complexity was first suggested by Chandra, Furst and Lipton. The best known lower bound, for an explicit function (in this model), is a lower bound of $\Omega(n/2^k)$, where n is the size of the input of each player, and k is the number of players (first proved by Babai, Nisan and Szegedy). In this note we extend the best lower bounds known to the best partition model.

This proof came from several conversations with Ran Raz and Avi Wigderson. After writing this note, we learned from Stasys Jukna that a similar construction was already obtained by Hayes [Hay01].

1 Introduction

The number-on-forehead (NOF) multiparty communication complexity model was first introduced by Chandra, Furst and Lipton [CFL83], as a generalization of Yao's standard 2-party communication complexity model [Yao79]. In the k-party number-on-forehead model, the players are all-powerful, but the input to f is partitioned into k pieces of n bits each, $x_1, \ldots, x_k \in \{0,1\}^n$, and x_i is placed, metaphorically, on the forehead of player i. Thus each player sees only $(k-1)n$ of the kn input bits. In order to compute f, the players communicate by writing bits on a shared blackboard, and the complexity of the protocol is the number of bits that are communicated (i.e., written on the blackboard). Since the introduction of the model in 1983, it has found applications in a large and diverse number of areas, including circuit complexity, pseudorandomness, and proof complexity [BNS92, NW93, Nis93, BPS05]. A major open problem is to prove a superpolylogarithmic lower bound for the k-party communication complexity of an explicit function, where k is superpolylogarithmic. By [BT91, HG91], this would prove superpolynomial lower bounds for the circuit class ACC, thereby solving a major open problem in circuit complexity. The best lower bound obtained so far is $\Omega(n/2^k)$, which breaks down when the number of players is logarithmic [BNS92, CT93, Raz00, FG06]. For an excellent survey of communication complexity, and multiparty communication complexity, see [KN97].

When computing a function f on kn bits in the multiparty model, we glossed over how the partition of the bits was chosen. By default, the partition is assumed to be *worst-case*. In this (worst-case) scenario, in order to prove a lower bound for f it suffices to find some partition such that f is hard in the NOF model for k players with respect to this partition. The best known lower bounds on the k-party communication complexity mentioned [BNS92] above are with respect to worst-case partitions.

In contrast, in the *best-case* partition model, in order to prove a lower bound for f, we must prove the lower bound for *any* partition of the kn bits into disjoint sets each of size n. Lower bounds in the best-case partition model have many applications, including VLSI layout, and arithmetic circuit complexity [RY09]. For $k = 2$ there are several papers showing how to prove linear lower bounds for best-partition partition from similar worst-case lower bounds [LR92, RY09]. However, in the number-on-forehead model for $k \geq 3$, there were no nontrivial lower bounds known with respect to best-case partitions. In this paper we give a reduction showing that for certain classes of functions, a worst-case partition lower bound for the function gives rise to a best-case lower bound for a related function. Using this result we obtain bounds similar to [BNS92] in the best-case setting, and we also separate some well-known NOF complexity classes with respect to best-case partitions.

2 Definitions

2.1 Multiparty Communication Complexity

In the k-party number-on-forehead communication complexity model, k players are trying to collaborate to compute a function $f : X_1 \times \ldots \times X_k \to \{0, 1\}$. For each player i, player i knows the values of all of the inputs $(x_1, \ldots, x_k) \in X_1 \times \ldots \times X_k$ except for x_i (which conceptually is thought of as being placed on Player i's forehead). The players exchange bits according to an agreed upon protocol, by writing them on a public blackboard. A *protocol* specifies what each player writes as a function of the blackboard content and the inputs seen by that player, and whether the protocol is over, in which case the last bit written is taken as the output of the protocol. The *cost* of a protocol is the maximum number of bits written on the blackboard.

In a *deterministic protocol*, the blackboard is initially empty. A *randomized protcol* is a distribution on deterministic protocols such that for every input, a protocol selected at random from the distribution, errs with probability at most $1/3$. In a *nondeterministic protocol*, an initial guess string is written on the blackboard at the beginning of the protocol (and counted towards the communication) and the players are trying to verify that the output of the function is 1 in the usual sense: there exists a string where the output of the protocol is 1 if and only if the output of the function is 1. The *communication complexity* of a function f under one of the above types of protocols is the minimum cost of a protocol of that type computing f.

In all communication models, a protocol is said to be *efficient* if it has complexity $\log n^{O(1)}$. Correspondingly, P_k^{cc}, RP_k^{cc}, BPP_k^{cc}, NP_k^{cc}, BPP_k^{cc} are the number-on-forehead analogs of the standard complexity classes [BFS86].

When computing a function f on kn bits in the multiparty model, we need to decide how the partition of the bits was chosen. By default, the partition is assumed to be *worst-case*. In this (worst-case) scenario, in order to prove a lower bound for f it suffices to find some partition such that f is hard in the NOF model for k players with respect to this partition. As mentioned earlier, best known lower bound on the k-party communication complexity of an explicit function f, for worst-case partitions, is $\Omega(n/2^k)$ [BNS92].

In contrast, in the *best-case* partition model, in order to prove a lower bound for f, we must prove the lower bound for *any* partition of the kn bits into disjoint sets each of size n. Analogous to the worst-case complexity classes, we can define the best-partition NOF communication complexity classes $\mathsf{P}_k^{cc,b}$, $\mathsf{RP}_k^{cc,b}$, $\mathsf{BPP}_k^{cc,b}$, $\mathsf{ZPP}_k^{cc,b}$, $\mathsf{NP}_k^{cc,b}$.

2.2 Hard Functions

Let G be a d-regular bipartite graph on $[n] \times [n]$ with normed adjacency matrix M. The largest eigenvalue of M is λ_1 with eigenvector $u = (1/n, \ldots, 1/n)$. The second largest eigenvalue is given by

$$\lambda_2 = \max_{||x||=1, x \perp u} ||Mx||.$$

The *spectral expansion* of G, denoted by $\lambda(G)$ is the value of λ_2.

Alon and Boppana [Alo86] proved that for sufficiently large n random d-regular bipartite graphs have spectral expansion at least $1/\sqrt{d}$.

Let x, y be boolean vectors, each of length n. The inner product of two vectors, IP(x, y) is equal to $\sum_i x_i y_i$ (mod 2). Let $x^1, \ldots, x^k$ be k vectors, each of length n. The generalized inner product of these vectors, GIP$(x^1, \ldots, x^k)$ is equal to $\sum_i x_i^1 \cdots x_i^k$ (mod 2). The set disjointness function, DISJ$(x^1, \ldots, x^k)$ is equal to $\vee_i (x_i^1 \wedge \ldots \wedge x_i^k)$.

Let G be d-regular bipartite graph with edges in $[n] \times [n]$. We will represent G by an n-by-n adjacency matrix A, where the entry $A(i, j)$ is 1 if and only if edge (i, j) is in G. Let Graph-IP$_A(x, y) = \sum_{i,j} A(i, j) x_i y_j$.

We can similarly define Graph-GIP by using a k-level graph as follows. Let H be a k-partite graph on $V_1 \cup \ldots \cup V_k$ where $|V_i| = n$ for all $i \leq k$. For each $i \leq k - 1$, assume that the subgraph from V_i to V_{i+1} is a degree d graph. We represent H by a k-dimensional tensor A where the entry $A(i_1, \ldots, i_k)$ is 1 if and only if for all levels l, $l \leq k - 1$, the edge (i_l, i_{l+1}) is in the l^{th} level of H. Let Graph-GIP$_A(x^1, \ldots, x^k) = \sum_{i_1,\ldots,i_k} A(i_1, \ldots, i_k) x_{i_1}^1, \ldots, x_{i_k}^k$. Similarly, let Graph-DISJ$_A(x^1, \ldots, x^k) = \vee_{i_1,\ldots,i_k} (A(i_1, \ldots, i_k) x_{i_1}^1, \ldots, x_{i_k}^k)$.

3 The Two Player Case

In this section we prove the following theorem.

Theorem 3.1. *Let G be a bipartite d-regular expander graph on vertices $V_1 \cup V_2$, with spectral expansion $\lambda = 1/\sqrt{d}$, and $d \geq 4$. If the best-partition (2-player) communication complexity of Graph-IP$_A(x, y)$ is at most $f(n)$, then the worst-case partition (2-player) communication complexity of IP(x, y) (over the standard partition) is at most $O(f(n))$.*

We will use the following well-known Expander Mixing Lemma [AC88].

Lemma 3.2. *(Expander Mixing Lemma) Let G be a d-regular bipartite expander from V_1 to V_2, $|V_1| = |V_2| = n$, with spectral expansion λ. Let $A \subseteq V_1$, $B \subseteq V_2$. Then the number of edges between A and B is at least $d|A||B|/n - \lambda d\sqrt{|A||B|}$. In particular when $|A| = |B|$, as long as $|A| \geq \lambda n$, there is at least one edge between A and B.*

Proof. (Proof of Theorem 3.1)

We want to show that if there exists a partition such that if Graph-IP is easy, then IP is also easy. Let $V = V_1 \cup V_2$ be the layers in the bipartite graph, $|V_1| = |V_2| = n$. Let P be a partition of $V_1 \cup V_2$ such that Graph-IP is easy with respect to this partition. P divides $V_1 \cup V_2$ into two classes, P_1 and P_2 where the variables associated with P_1 are given to Player 1, and the variables associated with P_2 are given to Player 2. Assume without loss of generality that P_1 contains a set $V_1' \subseteq V_1$ of at least $n/2$ vertices, and similarly P_2 contains a set $V_2' \subseteq V_2$ of at least $n/2$ vertices.

Our goal is to find large subsets $S \subseteq V_1'$, $T \subseteq V_2'$ where there is a matching between S and T and no other edges within S and T. We start with A_0 equal to V_1' and B_0 equal to V_2', where $|A_0| = |B_0| = n/2$. Choose some edge, $e = (i, j)$ from $A_0 \times B_0$, which exists by the Expander Mixing Lemma. Then remove all neighboring vertices of this edge and continue. The new set A_1 consists of all elements of A_0 except for the vertex i, and the $d-1$ other neighbors of j. Similarly B_1 consists of all elements of B_0 except for vertex j and the $d - 1$ neighbors of i. Thus $|A_1| = |B_1| = n/2 - d$.

By the Mixing Lemma, we can continue for s steps as long as $|A_s|, |B_s| \geq \lambda n = n/\sqrt{d}$. Thus we can continue as long as $n/2 - ds \geq n/\sqrt{d}$. Thus, s is linear in n for constant d. Let the edges of the matching be $e_1 = (a_1, b_1), e_2 = (a_2, b_2), \ldots e_s = (a_s, b_s)$.

Now consider an instance of IP where Player 1 has a string x of size s and Player 2 has a string y of size s. Player 1 creates an instance x' of Graph-IP on n bits with the above partition as follows. The bits of x are put in one-to-one correspondence with the variables $(a_1, a_2, \ldots a_s)$ that underly the set A_s, and all other bits of x' are set to zero. Similarly, Player 2 creates a string y' of length n by placing the variables of y in place of $(b_1, b_2, \ldots, b_s)$, and all other bits of y' are set to zero. It is clear that Graph-IP on x', y' computes the function IP on x, y. The size of x', y' is lineaer in the size of x, y and thus an upper bound of $f(n)$ for Graph-IP implies an upper bound of $O(f(n))$ for IP.

□

Note that the same lower bound works for set disjointness, or more generally for any function $f(x, y)$ where f computes a function $g(z_1, \ldots z_n)$ where $z_i = x_i \cdot y_i$. Also the reduction works in any model; if the original lower bound is randomized/nondeterministic/deterministic, then so is the best partition lower bound. Thus we obtain $O(n)$ randomized lower bounds in the best-partition model for Graph-IP, as well as for Graph-DISJ.

4 The General Case

In this section we will prove the following theorem.

Theorem 4.1. *Let G be a k-level graph with n vertices per level. Further assume that G has degree d and has spectral expansion $\lambda = 1/\sqrt{d}$, where $d = 16k^6$. There exists $\delta < 1$ such that for all $\epsilon < 1$, if for any $k < n^\delta$, the best-case k-party communication complexity of Graph-GIP$_A(x^1, \ldots, x^k)$ is at most n^ϵ, then the worst-case k-party communication complexity (for the standard partition) of GIP$(x^1, \ldots, x^k)$ is at most $n^{2\epsilon}$.*

Lemma 4.2. *Let $V = V_1 \cup V_2 \cup \ldots \cup V_k$, where each V_i is a set of n vertices. We think of V as the vertices of a levelled graph, with k levels, and n vertices per level. Let P be a partition of the vertices of V into k subsets, each of size n. Let P_i (called the i^{th} player) denote the elements of V that are given to the i^{th} player. Then there exists a one-to-one mapping $f : [k] \rightarrow [k]$ that assigns to each Player i a layer $V_{f(i)}$ such that for every $i \leq k$, the number of elements from $V_{f(i)}$ given to player i is at least n/l, where $l = 2k^2$.*

Proof. Fix a particular partition P. Based on the partition P, we create a k-by-k bipartite graph G as follows. The k left elements in G represent each of the k different V_i's. The k right elements in G correspond to the k different players (equivalence classes of the partition). The edge (i, j) is in G if and only if the number of vertices from V_i given to player P_j is at least n/l. It suffices to show that there exists a perfect matching in G. By Hall's Theorem, this is equivalent to proving that for every subset $C \subseteq [k]$ of vertices on the left, $|N(C)| \geq |C|$.

Fix some set C of size c, $1 \leq c \leq k$. Suppose for sake of contradiction that $|N(C)| < c$. Assume without loss of generality that the neighbors of C are $[1, \ldots, c-1]$. Since for every $j \geq c$, there is no edge from C to j, this means that for every $i \in C$, the number of elements from V_i given to Player j is at most n/l. Thus at least $n - (k - c + 1)n/l$ elements from V_i are given to players $[1, \ldots, c-1]$. Summing over all $i \in C$, this means that at least $c(n - (k - c + 1)n/l)$ elements from $\{V_i \mid i \in C\}$ are given to the set of players $P_1 \cup \ldots \cup P_{c-1}$. Since each player can receive at most n elements, this implies that $c(n - (k - c + 1)n/l \leq (c - 1)n$, which is a contradiction for $l = 2k^2$. □

Claim 4.3. *Let G be a k-partite graph on $V_1 \cup \ldots \cup V_k$, where $|V_i| = n$. Let $d = 16k^6$, $L = O(n/k^8)$. For each $i \leq k-1$, assume that the subgraph from V_i to V_{i+1} is a degree d graph with spectral expansion $\lambda = 1/\sqrt{d}$. Then there exist a layered subgraph G' on vertices $V'_1, V'_2, \ldots V'_k$, such that for all $i \leq k$, $|V'_i| = m$, $m = n/(2k^2)$, and such that there are L disjoint paths $p_1 \ldots p_l$ on G' and no other edges within G'.*

Proof. Start with a d-regular, layered graph, with n vertices per layer, and spectral expansion $\lambda = 1/\sqrt{d}$. We will first argue that there exists *one* path from $V_1 \cup \ldots \cup V_k$ that lies entirely within $V'_1 \cup \ldots \cup V'_k$.

Consider each bipartite subgraph G_i from V_i to V_{i+1}. By the Expander Mixing Lemma, for any subsets $S \subset V_i$, $T \subset V_{i+1}$, $|S| = |T|$, we are guaranteed to have at least one edge between S and T as long as $s_1 > n\lambda = n/\sqrt{d}$. Using the Expander Mixing Lemma repeatedly, we will argue that there must be a matching of size at least $(m - s_1)$ between V'_1 and V'_2. Starting with V'_1 and V'_2, find an edge which is guaranteed to exist by the Mixing Lemma as explained above. Remove the two vertices incident with this edge from V'_1 and V'_2 and find another edge, again using the Mixing Lemma. We can continue as long as the size of the set remaining is at least s_1. Thus we can find a matching between V'_1 and V'_2 of size at least $(m - s_1)$. Now let $V''_2 \subset V'_2$ be those vertices in V'_2 that are hit by the large matching that we just found. We will apply the same argument to find a matching between V''_2 and some subset V''_3 of size at least $(m - s_1) - s_1 = m - 2s_1$. After this step, we have found at least $m - 2s_1$ length 2 paths starting at V'_1, and lying entirely within $V'_1 \cup \ldots \cup V'_k$. Continuing the process k times, we will find at least $m - ks_1$ length k paths starting at V'_1 and lying entirely within $V'_1 \cup \ldots \cup V'_k$. As long as $(m - ks_1) \geq 1$, this will give us at least one path. For our choice of $m = n/(2k^2)$, this will hold as long as $k < 1/(2k^2\lambda) = \sqrt{d}/(2k^2)$.

Now we throw away all of the vertices along this path and their neighbors and continue in the same way to find another path. That is, let V'_i equal the old V'_i but with the vertices incident to this path removed, as well as all neighbors of this path. For each i, the size of the new V'_i will be at least $n/(2k^2) - 2d$. Repeat the argument above to find another length k path lying entirely within the new sets $V'_1 \cup \ldots \cup V'_k$.

We will continue this process L times, each time finding a length k path on $V'_1 \cup \ldots \cup V'_k$ that is vertex disjoint from all previously found length k paths. We can continue as long as $n/2k^2 - 2dL - ks_1 \geq 1$. Rearranging, we require $L < n/(4dk^2) - ks_1/(2d)$. Setting $d = 16k^6$, $s_1 = n/\sqrt{d}$, we require

$$\begin{aligned} L &< \frac{n}{4dk^2} - \frac{ks_1}{2d} \\ &= \frac{n}{64k^8} - \frac{kn}{2(16k^6)(4k^3)} \\ &= \frac{n}{64k^8} - \frac{n}{128k^8} \\ &= \frac{n}{128k^8} \end{aligned}$$

□

Proof. (Proof of Theorem 4.1)

As before, we want to show that if there exists a partition such that Graph-GIP is easy, then GIP is easy. Let $V = V_1 \cup \ldots \cup V_k$, $|V_i| = n$. Let P be a partition of V such that Graph-GIP is easy with respect to this partition. P divides V into k equal-sized classes, $P_1, \ldots, P_k$, where the variables associated with P_i are given to Player i. By Lemma 4.1, there exists an assignment f of levels to players such that for all i, $1 \leq i \leq k$, Player i is given at least $n/(2k^2)$ vertices from level $V_{f(i)}$. That is, $|P_i \cap V_{f(i)}| \geq n/(2k^2)$. For each j, let $W_j \subseteq V_j$ denote those vertices that are given to Player $f^{-1}(j)$. Assume without loss of generality that for each j, $|W_j| = m = n/(2k^2)$.

Now by Claim 4.1, there exists a subgraph G' on $\{W'_j \subseteq W_j \mid 1 \leq j \leq k\}$, such that $|W'_j| = s = O(n/k^8)$, and such that there are s disjoint paths $p_1, \ldots, p_s$ in G' and no other edges within G'.

Now consider an instance of GIP where player i has string x^i of size s. Given the above partition, player i creates an instance q^i of Graph-GIP on n bits as follows. The bits of x^i are put in one-to-one correspondence with the variables that underly the set W'_i and all other bits of q^i are set to zero. It is clear that Graph-GIP on q^i computes GIP on x^i.

Now suppose that Graph-IP on x', y' has an n^ϵ-bit communicion protocol. Using the above reduction and solving Graph-IP, gives us an n^ϵ-bit protocol for IP. In terms of s, this protocol has cost $(128k^8 s)^\epsilon$. Choosing k so that $128k^8 = n/(128k^8)$, the protocol has cost $(s)^{2\epsilon}$, as desired. When k is constant, we obtain linear best partition lower bounds from linear worst case partition bounds. For GIP, we have lower bounds of the form $n/4^k$ for all k. These translate to $n^{o(1)}$ best partition lower bounds for $k < \log n$.

□

4.1 Separations in the Best-case Model

As corollaries to Theorem 4.1, we obtain many separations of multiparty communication complexity classes in the best-case partition setting, as these follow from the worst-case separations and the observation that the function is of the type that we require for our reduction to go through.

Theorem 4.4. *There exists ϵ such that for all $k \geq \epsilon \log n$, $\mathsf{NP}_k^{cc} \neq \mathsf{BPP}_k^{cc}$*

References

[AC88] N. Alon and F. Chung. Explicit constructions of linear sized tolerant networks. *Discrete Math.*, 2:15–19, 1988.

[Alo86] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[BFS86] László Babai, Péter Frankl, and János Simon. Complexity classes in communication complexity theory (preliminary version). In *27th Annual Symposium on Foundations of Computer Science*, pages 337–347, Toronto, Ontario, Canada, 1986. IEEE.

[BNS92] László Babai, Noam Nisan, and Márió Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. Syst. Sci.*, 45(2):204–232, 1992.

[BPS05] Paul Beame, Toniann Pitassi, and Nathan Segerlind. Lower bounds for Lovasz-Schrijver systems and beyond follow from multiparty communication complexity. In *ICALP*, pages 1176–1188, 2005.

[BT91] Richard Beigel and Jun Tarui. On ACC. In *Proceedings of the 32nd Annual Symposium on Foundations of Computer Science*, pages 783–792, Los Alamitos, CA, USA, 1991. IEEE Computer Society Press.

[CFL83] Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *STOC '83: Proceedings of the fifteenth annual ACM symposium on Theory of computing*, pages 94–99, New York, NY, USA, 1983. ACM Press.

[CT93] Fan R. K. Chung and Prasad Tetali. Communication complexity and quasi randomness. *SIAM J. Discret. Math.*, 6(1):110–125, 1993.

[FG06] Jeff Ford and Anna Gál. Hadamard tensors and lower bounds on multiparty communication complexity. In *Complexity of Boolean Functions*, 2006.

[Hay01] Thomas Hayes. Separating the k-party communication hierarchy: an application of the zarankiewicz problem. In *Unpublished manuscript*, 2001.

[HG91] J. Håstad and M. Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1(2):113–129, 1991.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, New York, NY, USA, 1997.

[LR92] T. W. Lam and W. L. Ruzzo. Results on communication complexity classes. *J. Computer Systems Sciences*, 44:324–342, 1992.

[Nis93] Noam Nisan. The communication complexity of threshold gates. In *Proc. of "Combinatorics, Paul Erdos is Eighty"*, pages 301–315, 1993.

[NW93] Noam Nisan and Avi Wigderson. Rounds in communication complexity revisited. *SIAM J. Comput.*, 22(1):211–219, 1993.

[Raz00] Ran Raz. The BNS-Chung criterion for multi-party communication complexity. *Comput. Complex.*, 9(2):113–122, 2000.

[RY09] R. Raz and A. Yehudayoff. Multilinear formulas, maximal partition discrepancy, and mixed source extractors. In *Proceedings 49th IEEE FOCS*, 2009.

[Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing(preliminary report). In *STOC '79: Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213, New York, NY, USA, 1979. ACM Press.